Biomarkers of insulin resistance and their performance as predictors of treatment response in adults with risk factors for type 2 diabetes.

RJ Brogan[1,2], O Rooyackers[3], BE Phillips[4], B Twelkmeyer[3], LM Ross[5], PJ Atherton[3], WE Kraus[5], JA Timmons[2,6*] and IJ Gallagher[7*]

Affiliations

[1]Anaesthesia, Pain and Perioperative Medicine, Fiona Stanley Hospital, Perth, WA, Australia
[2]Augur Precision Medicine LTD, Stirling, Scotland
[3]Division of Anesthesiology and Intensive Care, CLINTEC, Karolinska Institutet, Sweden
[4]Clinical, Metabolic and Molecular Physiology Research Group, School of Medicine, University of Nottingham, Derby, England
[5]Duke University School of Medicine, Durham, NC, USA
[6]Faculty of Medicine and Dentistry, Queen Mary University of London, London, England
[7]Center for Biomedicine and Global Health, Edinburgh Napier University, Edinburgh, Scotland
* Senior Authors

Key words: T2DM, Insulin resistance, biomarkers, regression, Bayesian projective prediction

Corresponding author: robert.brogan@health.wa.gov.au

Abstract

Insulin Resistance (IR) is a component of the pathogenesis of type 2 diabetes mellitus (T2DM), and risk factor for cardiovascular and neurodegenerative diseases. Amino acid and lipid metabolomic IR diagnostics associate with future T2DM risk in epidemiological cohorts. Whether these assays can accurately detect altered IR following treatment has not been established. In the present study we evaluated the ability of metabolomic diagnostics to predict altered IR following exercise treatment. We evaluated the performance of two distinct insulin assays and built combined clinical and metabolomic IR diagnostics. These were utilised to stratify IR status in the pre-intervention fasting samples in three independent cohorts (META-PREDICT (MP, n=179), STRRIDE-AT/RT (S-2, n=116) and STRRIDE-PD (S-PD, n=149)). Linear and Bayesian projective prediction strategies were used to evaluate biomarkers for fasting insulin and HOMA2-IR and change in fasting insulin with treatment. Both insulin assays accurately quantified international standard insulin (R^2>0.99), yet agreement for fasting insulin was less congruent (R^2=0.65). Only the high-sensitivity ELISA assay could identify the mean effect of treatment on fasting insulin. Clinical-metabolomic models were statistically related to fasting insulin (R^2 0.33 – 0.39) but had modest capacity to classify HOMA2-IR at a clinically relevant threshold. Furthermore, no model predicted treatment responses in any cohort. Thus, we demonstrate that the choice of insulin assay is critical when quantifying the influence of life-style on fasting insulin, while none of the clinical-metabolomic biomarkers, validated in cross-sectional data, are suitable for monitoring longitudinally changes in IR status.

50 **Introduction**

51

52 Chronic hyperinsulinaemia, a feature of type 2 diabetes mellitus (T2DM), contributes to a cycle of
53 events driving further insulin resistance (IR) [1], promoting cardiovascular [2] and neurodegenerative
54 diseases [3]. Extreme IR also represents a distinct category of T2DM [4]. Lifestyle based treatments
55 are effective at reducing the risk of developing T2DM. [5–8] Such programs, typically involving
56 exercise training, modify IR with high intra-individual variability. [9–14] Similar variability is
57 observed in response to drug treatment. [15,16] While monitoring the efficacy of treatment on glucose
58 tolerance in large cohorts is relatively straightforward [17], the same cannot be said for monitoring
59 insulin action. [18,19] Furthermore, the varying characteristics of commonly used insulin assays
60 complicate comparison across studies. [20,21] There are clear clinical definitions for diagnosing
61 T2DM and impaired glucose tolerance (IGT) but less so for IR, possibly due to the lack of
62 widespread measurement of insulin. This is problematic because IR is distinct from glucose
63 intolerance and can precede T2DM and slowly developing age-related diseases by decades and
64 thus there is an urgent need to enable tracking over time IR. [22,23]

65

66 This need for cost-effective practical IR diagnostics has stimulated the development of several
67 clinical [24] and molecular prototypes, including assays relying on branched-chain amino acids
68 (BCAA)[25–28], plasma lipoproteins[29–31] and multi-protein signatures.[32] The abundance of several
69 circulating amino acid and plasma lipids have been incorporated into statistical models that predict
70 IR in cross-sectional, prospective or retrospective analyses (i.e. predict T2DM risk).[25–27,29–31,33]
71 Circulating BCAA are particularly attractive as surrogates for IR because disrupted catabolism of
72 individual BCAA have been mechanistically and causally linked to IR, cardiovascular and

73 neurodegenerative disease processes in recent years [13,28,34–43]. It remains unclear how well
74 abundance of BCAA in the circulation reflects the mechanistic role of BCAA within the key tissues
75 responsible for their catabolism and therefore in turn their ability to track with changes in
76 metabolic fitness.
77
78 In addition to predicting IR, it would therefore be ideal if metabolic 'risk' biomarkers could
79 reliably associate with improvements in IR following treatment, so that non-responders could be
80 reallocated to alternative treatments. [14–16,44,45] There have been attempts to establish if plasma
81 BCAA abundance tracks treatment responses yielding conflicting conclusions, partly because
82 existing studies have been too small to reliably explore such relationships. [46,47] Determinants of
83 insulin action also reflect a combination of distinct acute mechanisms in the hours post-exercise,
84 and more stable long-term adaptive changes [48] – such as increased tissue vascularisation. [49] Such
85 temporal influences have not been considered when evaluating these novel biomarkers.
86 Immediately post-exercise insulin action and glucose tolerance can be impaired, while insulin
87 responses to an oral glucose tolerance test (OGTT) appear lowest 72hr following the previous
88 exercise training session.[50]
89
90 In the present analyses we have assessments of IR within 24hr and 48-72hr after the last training
91 session of chronic exercise training that was effective in reducing group mean HOMA2IR. The
92 idea is that novel biomarkers should offer practical (e.g. cost, simplicity) advantages over
93 laboratory insulin assays, and if they can be shown to predict longitudinal changes in IR, then
94 efforts to configure low-cost standardised assay from minimally invasive sampling would be
95 motivated. In the present multi-study analysis, we evaluated, in hundreds of individuals at-risk for

96 T2DM, the performance of established BCAA and lipid biomarkers, combined with simple to
97 measure clinical phenotypes to estimate fasting IR and predict changes in insulin following
98 supervised lifestyle interventions.[9,12,51,52] We established that choice of insulin assay is critical for
99 capturing interactions between exercise status and fasting insulin, while plasma biomarkers of IR
100 status, from cross-sectional and prospective studies, are unable to track with treatment responses.
101

102 **Results**

103

104 Demographic, blood and metabolomic data for the three independent cohorts - STRRIDE AT/RT
105 (S-2), STRRIDE-PD (S-PD) and MP - are presented in **Table 1.**

106

107 *Comparison of insulin assay and metabolomics assay performance*

108 A long-standing challenge is that commercial assays for insulin show distinct specificity and
109 sensitivity profiles, with no agreement to move to a single standard assay.[20,21] Some historical
110 insulin assays, utilised in exercise intervention studies (e.g. HERITAGE)[53] or included in genome-
111 wide association modelling,[54] do not show molecular associations consistent with modern more
112 specific insulin assays.[13] We produced a large-scale comparison of two typical insulin assays (the
113 immunoassay based automated analyzer (Immulite 2000) and a manual high sensitivity ELISA kit
114 (Dako, Stockholm, Sweden), run in parallel on samples obtained before and after exercise-training
115 that reduced IR.[55] Both Immulite 2000 and Dako ELISA showed near perfect ($>R^2=0.99$)
116 correlation with the international WHO standard for insulin (National Institute for Biological
117 Standards and Control, **Figure S1A-B**). No cross reaction with c-peptide standard was detected
118 (data not shown). When plotting insulin values from the same fasting samples, there was moderate

agreement between the two assays (R^2=0.65, **Figure S2)**. Critically, insulin measured using the high-sensitivity ELISA was able to detect a group mean reduction following exercise training, while the Immulite 2000 immune-assay could not (**Figure 1**). All the subsequent analyses relied on the data generated by the Dako high-sensitivity ELISA (produced in a single laboratory).

Biomarker association with fasting insulin and HOMA2-IR status

Predictor variables were selected based on the literature summarised above, and availability in all three cohorts. Linear regression was performed to explore individual relationships between fasting insulin and the selected clinical and metabolomic biomarkers (**Figure 2**; full results in **Table S1**). In MP, all pre-selected variables other than age (p=0.09) reached statistical significance for association with fasting insulin (age is a reliable covariate but the age-range in MP was limited). In S-2 $\log_{10}$ BMI (p<0.001), fasting glucose (p<0.001), HDL (p=0.007), TAG (p<0.001) and amino acids (isoleucine (ILE) p=0.0005, leucine (LEU) p=0.006, valine (VAL) p<0.001, alanine (ALA) p=0.0027) were linearly associated with fasting insulin. Similar levels of associations were observed in the S-PD cohort, except for ALA (p=0.068). For multi-variable predictors of fasting insulin (or HOMA2-IR) and predictors of treatment responses, variables which showed a statistically significant association in one or more of the cohorts were considered (as each was supported with published evidence for having an association in cross-sectional cohorts).

As each study used some distinct metabolomic methods (see Methods) we performed multivariable linear regression and logistic regression with K-fold cross validation (k=10) separately in MP, S-2 and S-PD cohorts. Four models were considered; a baseline model (age, BMI and gender), a BCAA model (age, BMI, gender and sum of BCAA), a lipid model (age, BMI, gender, HDL, LDL

142 and TAG) and finally, a combined model (age, BMI, gender, BCAA, HDL, LDL and TAG). The
143 combined model estimated fasting insulin (R^2 0.33 – 0.39), with the results in agreement with
144 previous observations (**Table 2**). Fasting glucose was excluded when predicting IR status. For IR
145 ROC analysis, insulin sensitive (IS, HOMA2-IR <1.3) and insulin resistant (IR, HOMA2-IR ≥1.3)
146 was defined based on results from a large population study (n=95,540) where HOMA2-IR ≥1.3
147 was associated with an increased risk for the development of T2DM over a median of 4.7y (hazard
148 ratio of 3.2 (95% CI 1.9-5.3)). [56] The four models were assessed in each cohort for their ability to
149 classify IS vs IR (**Figure 3**). The baseline model had no discriminatory performance in S-2, so the
150 metabolomics variables improved that model substantially, however metabolomics did not add
151 further to the clinical model ROC performance in MP or S-PD, and the best models had a modest
152 Cohens Kappa (**Table S2)**. This illustrates that even though BCAA and lipid models statistically
153 associate with IR status, they do not contribute to a clinically useful model.
154
155 *Biomarker based prediction of IR status in response to lifestyle intervention.*
156 We established that each biomarker statistically associated with fasting insulin in each study
157 (**Table S1**) and as reported by others. Finally, we examined the utility of these biomarkers for
158 *predicting* exercise treatment related insulin responses in IR. For this analysis we were able to
159 include changes in fasting glucose (easy and inexpensive to measure) as we did not rely on the
160 HOMA model. **Figure 4** displays the univariable linear relationships of changes in clinical and
161 metabolomic biomarkers with changes in fasting insulin as measured with the high sensitivity
162 ELISA over an exercise intervention in all three cohorts (**Figure S3** provides the correlation
163 coefficients (including for individual amino acids). To examine the classification utility of
164 individual and combinations of biomarkers we built Bayesian linear models, independently for

each cohort, modelling change in fasting insulin after each intervention. The models included age and changes over the intervention in individual BCAA, fasting glucose, BMI, lipids, systolic blood pressure and alanine. The full models ("reference" models) were considered as the optimum solution to the prediction task.[57] The posterior predictive distributions showed the models performed well and were able to predict data like that of the observed data (**Fig 4A-C**). Using projective prediction, we found that the biomarkers did not, however, improve predictive performance over a null model i.e. the intercept term (**Figure 4D-F**).[58] Thus, change in biomarkers (alone or additively combined) could not predict change in fasting insulin following an exercise intervention.

Discussion

Development of a cost-effective diagnostic for insulin sensitivity would help better understand the aetiology of T2DM and aligning clinical practice with the key recommendations of the International Diabetes Federation, namely early detection and prevention of pre-diabetes.[59] Hyperinsulinaemia drives altered insulin receptor expression potentially leading to further IR.[1] IR in turn contributes to cardiovascular disease[2] and dementia risk.[3] Insulin is not, however, routinely measured in the clinic and thus crucial information regarding the temporal nature of IR is lacking. Previous work has identified several potential biomarkers for IR in cross-sectional and prospective settings.[25–27,29–31,33] The present study extends this body of work to evaluate these IR biomarkers in the context of their ability to accurately classify changes in fasting insulin in response to exercise (MP and S2) or lifestyle (S-PD) based treatment for individuals with multiple risk-factors for T2DM.

188 *Importance of the insulin assay*

189 To our knowledge we are the first to demonstrate the importance of insulin assay choice on
190 detecting changes in fasting insulin in large scale treatment studies. The MP cohort represents a
191 typical group of sedentary middle-aged individuals with multiple risk factors for T2DM.[12] We note
192 that the Immulite 2000 immunoassay was unable to detect exercise induced improvements in
193 fasting insulin, while concurrent measurements using a high sensitivity ELISA assay did. Notably
194 neither assay demonstrated any cross-reactivity with c-peptide and they both performed well in a
195 standard curve analysis. The American Diabetes Association attempted to reach a consensus on
196 standardizing the measurement of insulin.[20,21] Our analysis indicates that if environmental
197 influences or treatment effects on fasting insulin levels are to be captured, a high-sensitivity ELISA
198 assay should be utilised. We have previously reported that some older insulin assays, with
199 undocumented specificity (e.g. HERITAGE),[60] were unsuitable for genomic association
200 analysis.[13] Indeed, it is plausible that the use of insulin assays insensitive to variations across the
201 'normal' range of fasting insulin, or with cross-reactivity for C-peptide, have undermined genome-
202 wide association modelling of fasting insulin.[61] While suboptimal insulin assays will not be
203 voluntarily retired by manufacturers, it would seem sensible that the research community adopt
204 the use of high-sensitivity assays for studies of fasting insulin in all situations (other than extreme
205 hyperinsulinaemia)[20,21] especially when subtle gene-environment interactions are of interest to
206 study.

207

208 *Metabolomic components of the IR biomarker models*

209 There is a well-established positive correlation between individual circulating BCAA and greater
210 IR [25–27] and pathways that control BCAA catabolism have gained interest as possible targets for

211 reversing IR.[28,62,63] Furthermore, BCAA have been combined with clinical parameters to
212 reportedly increase the performance of models that predict both IR and risk of T2DM.[29,31,33] We
213 previously identified, using four independent cohorts, a transcriptional signature including BCAA
214 catabolism pathway genes, that related to IR after adjustment for age, VO2max and BMI (n=564).
215 We also found that in response to four independent lifestyle interventions (n=196), expression of
216 16 genes in human skeletal muscle changed in proportion to improvements in insulin sensitivity,
217 with 25% of those related to BCAA metabolism.[13] While this makes sense, as by mass, skeletal
218 muscle has the largest capacity for BCAA catabolism in humans [64] it does not provide evidence
219 that circulating levels of BCAA's will dynamically reflect tissue responses to insulin. In fact,
220 insulin mediated BCAA clearance may also rely on other organs including the liver.[65]
221
222 In a recent study, Lee *et al* claimed that baseline plasma BCAA levels related to change in glucose
223 infusion rate during a hyperinsulinaemic−euglycaemic clamp (a measure of insulin sensitivity and
224 liver fat content [47]). This was despite no change in total plasma BCAA levels and an increase in
225 muscle BCAA catabolism pathway gene expression. We also found no change in group mean
226 blood BCAA abundance with any exercise training protocol (**Table 1**), which might be
227 understandable as only ~40% of subjects demonstrate a numerical improvement in HOMA2-IR
228 and some demonstrate an increase in IR.[12,13] Nevertheless we did not find that changes in combined
229 plasma BCAA were able to predict changes in IR status. This observation is consistent with earlier
230 modelling approaches applied to the STRRIDE cohorts, where the combination of BCAA and
231 lipid-based IR biomarkers could not predict changes in a measure of insulin sensitivity.[17] Various
232 other combinations of BCAA and protein metabolites have been considered as IR models and
233 while some shared variance has been observed,[66] these are most likely attributable to random

234 differences in small cohorts. Using energy restriction to modify IR, we recently reported that group
235 mean differences in plasma BCAA abundance were also unchanged, despite group mean
236 improvements in IR. [66] All of these observations contrast with a previous study, relying on a sample
237 size a 10th of the present work, where changes in plasma BCAA abundance did track with improved
238 IR status.[46]
239
240 We attempted several modelling strategies, including several distinct clinical and metabolomic
241 variables, and modelled both fasting insulin and HOMA2-IR values to ensure our negative
242 conclusions were rigorous. We used Bayesian modelling and projective prediction [67,68] to establish
243 a sparse model that retained the ability to provide useful predictions of circulating insulin changes.
244 Projective prediction uses the predictions from the reference (full) model to identify a subset of
245 variables with predictive ability as close as possible to that of the full model. We were unable to
246 identify any set of variables useful for predicting change in circulating insulin after lifestyle
247 intervention using this strategy. Conventional regression modelling, against changes in HOMA2-
248 IR also failed to identify a useful model. We can conclude that the present clinical-metabolic
249 measures are not useful for classifying a HOMA2-IR threshold value indicative of 3-fold elevated
250 risk of T2DM over 5 years, nor track altered fasting insulin with life-style interventions. This is
251 even though our ELISA-based insulin values were strongly associated with causal molecular
252 drivers of T2DM and insulin signalling in tissue samples.[13]
253
254 *Clinical evaluation of IR biomarker models*
255 Diagnosing IR is challenging not least because defining IR remains controversial. The gold
256 standard method is commonly described as the 'clamp' technique or intravenous glucose tolerance

257 test.[19,69] However, substantial resources are required for this such that it is rarely used for diagnosis
258 and by their nature both can be supraphysiological (re levels of insulin). Alternative IR indices
259 exist and include the homeostatic model assessment (HOMA and HOMA2 [70,71], Matsuda index [72]
260 and the disposition index (DI).[73] These are not used in clinical practice, largely due to a lack of
261 routine measurement of insulin. Although several epidemiological studies use HOMA and
262 HOMA2 for modelling IR, there is limited basis for defining a clinical cut off value. Several studies
263 unwisely use a percentile cut off identified within their individual cohort, and then apply ROC
264 curve analysis to determine the threshold for a diagnosis of IR and then relate this to relative risk
265 for relevant clinical outcomes e.g. progression to T2DM or cardiovascular disease. As a result, the
266 HOMA IR threshold used across studies, to define 'at risk', has varied from 1.8 to 3.9 depending
267 on the population studied and methods used.[18] This represents a form of model overfitting.
268
269 It was estimated that a HOMA2-IR more than 1.3 represented a hazard ratio of >3.2 for developing
270 T2DM over ~5 years in 95,450 subjects.[74] This analysis is the largest study to date linking a
271 particular HOMA2-IR value to incipient diabetes and we used this cutoff in the present study.
272 Interestingly, in exploratory analysis, 83% of those with the most severe IR (HOMA2-IR≥2.4)
273 were not classified as metabolically compromised when using fasting glucose or their 2-hour-
274 glucose OGTT data.[12] The current reliance on glucose-centric risk assessment also limits our
275 ability to identify and hence intervene in those people with the "least effective" response to insulin
276 (defined as the totality of all of insulin's key physiology). Nevertheless, all the surrogate variables
277 investigated in the present study have limited ability to *diagnose* IR in a cross-sectional setting,
278 and *no* ability to track changes in IR with life-style treatment.

279 In summary, we evaluated multiple predictive models for IR, including models incorporating well
280 validated cholesterol species and BCAA. These models failed to forecast change in IR status
281 following supervised lifestyle modification. A limitation of our study is the limited metabolomic
282 variables gathered across cohorts (although this would mainly impact on the identification of novel
283 metabolites to consider). Nevertheless, any new blood based metabolic disease biomarkers
284 reported to track improvements in insulin sensitivity across multiple independent cohorts would
285 require substantial independent clinical studies to validate beyond the level of the models evaluated
286 in the present study.
287
288
289 **Materials and Methods**
290
291 *Clinical Cohorts*
292 The MP cohort consisted of 189 active participants, recruited as previously described from five
293 geographical regions across Europe.[12] All clinical methods relied on cross-site standard operating
294 procedures (SOPs), while insulin and metabolomic measures were performed in a single central
295 laboratory. All participants were classified as sedentary (<600 metabolic equivalents (METs)
296 $min \cdot wk^{-1}$) using a modified International Physical Activity Questionnaire,[75] and had a fasting
297 blood glucose level consistent with World Health Organisation (WHO) criteria for impaired
298 glucose tolerance ($5.5 < IGT < 7.0\ mmol \cdot l^{-1}$), and/or a BMI >27 $kg \cdot m^{-2}$. Participants were excluded
299 if they had evidence of active cardiovascular, cerebrovascular, respiratory, gastrointestinal or renal
300 disease or had a history of malignancy, coagulation dysfunction, musculoskeletal or neurological
301 disorders, recent steroid or hormone replacement therapy, or any condition requiring long-term

302 drug prescriptions. Exercise training was supervised and consisted of three high intensity cycling
303 sessions per week for 6 weeks as previously described.[12] Prior to the baseline assessments,
304 participants were instructed to refrain from exercise for three days. Supine blood pressure (Omron
305 M2, Omron Healthcare, Kyoto, Japan) and resting heart rate (RHR) were determined as the
306 average of three consecutive measurements. Blood glucose (from a dorsal hand vein) was analysed
307 before and during an (OGTT) using a YSI 2300 STAT Plus glucose analyser (Yellow Springs Inc.
308 Ohio, USA). Blood was collected in lithium heparin spray-coated vacutainers (Becton Dickinson,
309 New Jersey, USA) for analysis of insulin. The MP trial registration code is NCT01920659. The
310 STRRIDE cohorts were composed of sedentary middle-aged individuals at risk of developing
311 cardio-metabolic disease. [9,51,52] Following baseline assessments, each subject was randomized to
312 one of a variety of exercise training programs for six to eight months with or without a modified
313 diet.[9,51,52] Pre and post clinical and physiological assessments and outcomes were as previously
314 reported.[9,51,52]
315
316 *Plasma sample insulin analysis*
317 The samples from MP and the STRIDDE studies were analysed using high sensitivity insulin
318 ELISA (Dako A/S, Sweden). Two levels of QC solutions were run for insulin. Coefficients of
319 variation (CV) were acceptable; 4.68-8.03 % on both levels. As we were also using the
320 automated analyser (Immulite 2000) for analyses of all OGTT time points for insulin and c-
321 peptide from the MP study, we also assessed available fasting sample profiles using this lower
322 sensitivity immune-assay. To assure comparability of the results, we independently checked the
323 assay performance for cross detection of insulin and C-peptide. The WHO standard for insulin
324 and C-peptide were obtained from NIBSC (The National Institute for Biological Standards and

325 Control). A dilution series for both analytes was prepared and run on Dako ELISA and Immulite
326 2000. The diluted WHO insulin standards were subjected to the same procedure as the samples.
327 The results for both Immulite 2000 and Dako ELISA showed an acceptable correlation of
328 expected versus measured concentration with $R^2 = 0.9995$ and $R^2 = 0.9923$ for Immulite 2000
329 and Dako ELISA, respectively. The results are presented in **Figure S1**. No cross reaction
330 between insulin and c-peptide were detected. The two assays were in broad agreement over the
331 large concentration range of insulin used in this standard assay evaluation protocol **Figure S2**.
332
333 *Plasma sample metabolomic analysis*
334 Metabolomic analyses for MP has not been previously published (other than fasting glucose and
335 insulin). All samples were randomized for sample preparation and analyses resulting in an equal
336 distribution of all centres and possible responders and non-responders. We used hierarchical
337 clustering of metabolomic variables to highlight that there was no systematic centre specific bias
338 (**Figure S4**). All samples were taken after an overnight fast as arterialized venous blood and run
339 in duplicate. Results within ±20% CV were accepted unless otherwise stated; if this limit was not
340 met the sample was rerun. Quality controls (QC) were run for all analyses and with each batch of
341 sample preparation to ensure stable performance of the analysis procedure. QC for all analyses
342 included in house controls made from pooled EDTA plasma from healthy controls originating
343 from a previous study. The following methods were used for each metabolite. Sample analysis
344 used a Konelab 20XTi photospectrometer (Thermo Fisher Scientific, Thermo Electron Oy, Vantaa
345 Finland) for high density lipoprotein HDL cholesterol, LDL cholesterol and triacylglycerides
346 (TAG). Amino acids were analyzed using HPLC 2695 (Waters, Watford, UK) with online
347 derivatization and a fluorescence detector 474 (Waters, Watford, UK) using a method described

348 previously.[76,77] Fatty acids were analyzed using ultra-high performance liquid chromatography
349 (Ultimate 3000 UHPLC, Thermo Fisher Scientific, Germering, Germany) coupled to a TSQ
350 Vantage with electrospray ionization (ESI) and triple quadrupole mass spectrometer (MS/MS)
351 (Thermo Fisher Scientific, San Jose, CS, USA). For the blood obtained from the STRRIDE
352 cohorts' samples were analysed for plasma glucose concentration at the research site (Duke
353 University, USA) using a Beckman–Coulter DxC600 clinical analyser (Brea, CA, United States).
354 Insulin was analyzed (in the Stockholm lab) using the high sensitivity insulin ELISA assay (Dako
355 Denmark A/S, Glostrup, Denmark); and fasted plasma samples were analysed on 400 MHz nuclear
356 magnetic resonance profilers at LipoScience, now LabCorp (Morrisville, NC, United States), as
357 previously described [78]. The lipoprotein parameters and the BCAA were identified by
358 retrospectively analysing digitally stored spectra using the newly developed NMR-based
359 lipoprotein LP4 algorithm, which correlate with mass spectrometry methods.[79]
360
361 As distinct methodologies were utilised for the metabolomic analyses between MP and STRRIDE
362 cohorts we explore the compatibility of the common metabolomic data used in our diagnostic
363 models. For S-2 and S-PD, two methodologies were utilised for lipid related metabolomics and
364 Bland-Altman analysis was performed to assess measurement agreement between HDL and TAG
365 with NMR versus mass spectroscopy (**Figure S5-8**). There was good linear agreement between
366 methods for HDL (S-2 r^2 0.96, $p<0.001$; S-PD r^2 0.89, $p<0.001$) and TAG (S-2 r^2 0.99, $p<0.001$;
367 S-PD r^2 0.95, $p<0.001$). Bland-Altman analysis revealed a mean measurement error of -0.06
368 mmol/L for HDL (S-2 **Figure 5B**, S-PD **Figure S7B**) and +0.04mmol/L (S-2 **Figure 6B**, S-PD
369 **Figure S8B**) for TAG with 95% of observations within 1.96 standard deviation of mean. Work by
370 Wolak-Dinsmore *et al* demonstrated that NMR overestimates VAL and LEU by >25% (mean

371 value) and underestimates ILE by 10-15% (mean value) and while MS was used in MP, S-2 and
372 S-PD relied on NMR for BCAA quantification. Based on these known systematic differences,[79]
373 modelling was performed independently in each cohort using cross-validation methods.
374
375 *Statistical Modelling*
376 Raw data from the clinical chemistry analyses was analysed with GraphPad Prism 5 (Software
377 MacKiev, 2007, version 5.01) and more advanced analysis was accomplished using STATISTICA
378 10 (StatSoft Inc., 2011, version 10.0.228.2). All subsequent analysis was performed in R version
379 4.1.1 and 4.3.2. We utilised conventional and Bayesian strategies to model data in the present
380 work. Potentially predictive variables were selected based on known associations with IR and
381 T2DM. Linear regression modelling was applied to investigate associations between dependent
382 and independent variables. Briefly, after reducing each dataset to complete cases common
383 variables were identified across all three cohorts. The criterion for considering a clinically useful
384 variable was pragmatic – it had to be something that could be measured cheaply and reliably in a
385 primary care setting. Laboratory measures like maximal aerobic capacity related to IR status in
386 some but not all cohorts but was not considered because of the time and cost required to measure
387 it accurately. Several variables were selected based on these ease/low-cost criteria; some variables
388 were dropped because of high collinearity (e.g. systolic blood pressure was retained in favour of
389 diastolic blood pressure and mean arterial pressure in the Bayesian analysis). We removed
390 participants with obesity class III (BMI >40kg/m^2) for the Bayesian analysis and all continuous
391 variables were normalised. The latter steps were carried out using the tidymodels R package
392 (10.32614/CRAN.package.tidymodels) and applied consistently across all datasets. Collinearity
393 between variables was checked for each dataset separately **(Figure S3)**. Apart from BCAA in the

394 MP dataset there were no pairwise correlations with an absolute value >0.5. Linear models were
395 built with student-t priors on both the intercept and slope coefficient terms (df=3, mean=0 and
396 scale=4). These were mildly informative on the scale of the normalised data. Details are in the
397 supplementary material.
398
399 We built Bayesian multivariable linear models using the brms package to predict change in
400 circulating insulin using relevant clinical variables, circulating BCAA levels and selective lipids.[80]
401 These full models were considered reference models. We then used projective prediction to
402 identify a subset of variables with predictive performance as close as possible to that of the
403 reference model.[58] Briefly projective prediction first generates a solution path - the variable
404 ranking - for each sub model examined. Next a leave-one-out cross-validation process determines
405 the predictive performance of each sub model along the predictor ranking.
406
407 **Funding**
408 The clinical data utilised in this study was funded by multiple sources: European Union Seventh
409 Framework Programme (META-PREDICT, HEALTH-F2-2012-277936); STRRIDE II
410 (NCT00275145) by NHLBI grant HL-057354 and STRRIDE-PD (NCT00962962) by NIDDK
411 DK-081559 and R01DK081559. The data modelling was supported by Augur Precision Medicine
412 LTD.
413
414 **Competing interests**
415 None declared. JT is a major shareholder in APM, however there is no commercial link between
416 APM and the present study. There are no other financial conflicts of interest to disclose.

Contributions

RJB, JAT, OR and IJG conceptualized aims and designed the analysis strategy. All authors contributed to the project design, data collection and/or pre-processing. RJB and IJG performed data analysis and JAT contributed to interpretation. JAT and RJB wrote the first draft and JAT, RJB and IJG produced a full manuscript that all authors contributed to and have read and approved the final manuscript. JT wrote the HEALTH-F2-2012-277936 grant application.

Data and code availability

Some or all datasets generated during and/or analysed during the current study are not publicly available but are available from the corresponding author on reasonable request. Data and the analysis script can be found at **10.5281/zenodo.13819578.**

References

1. Cen HH, Hussein B, Botezelli JD, et al. Human and mouse muscle transcriptomic analyses identify insulin receptor mRNA downregulation in hyperinsulinemia-associated insulin resistance. *FASEB Journal*. 2022;36(1). doi:10.1096/fj.202100497RR

2. Wamil M, Coleman RL, Adler AI, McMurray JJV, Holman RR. Increased Risk of Incident Heart Failure and Death Is Associated With Insulin Resistance in People With Newly Diagnosed Type 2 Diabetes: UKPDS 89. *Diabetes Care*. 2021;44(8):1877-1884. doi:10.2337/DC21-0429

3. Folch J, Olloquequi J, Ettcheto M, et al. The Involvement of Peripheral and Brain Insulin Resistance in Late Onset Alzheimer's Dementia. *Front Aging Neurosci*. 2019;11(September):1-16. doi:10.3389/fnagi.2019.00236

4. Dennis JM, Shields BM, Henley WE, Jones AG, Hattersley AT. Disease progression and treatment response in data-driven subgroups of type 2 diabetes compared with models based on simple clinical features: an analysis using clinical trial data. *Lancet Diabetes Endocrinol*. 2019;7(6):442-451. doi:10.1016/S2213-8587(19)30087-7

5. Knowler WC, Fowler SE, Hamman RF, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. *Lancet*. 2009;374(9702):1677-1686.

446 6. Wing RR, Bolin P, Brancati FL, et al. Cardiovascular effects of intensive lifestyle intervention in
447 type 2 diabetes. *New England Journal of Medicine*. 2013;369(2):145-154.
448 doi:10.1056/NEJMoa1212914

449 7. Uusitupa M, Peltonen M, Lindström J, et al. Ten-year mortality and cardiovascular morbidity in
450 the Finnish Diabetes Prevention Study - Secondary analysis of the randomized trial. *PLoS One*.
451 2009;4(5):1-8. doi:10.1371/journal.pone.0005656

452 8. Hivert MF, Christophi CA, Franks PW, et al. Lifestyle and metformin ameliorate insulin
453 sensitivity independently of the genetic burden of established insulin resistance variants in
454 diabetes prevention program participants. *Diabetes*. 2016;65(2):520-526. doi:10.2337/db15-0950

455 9. Ross LM, Slentz CA, Kraus WE. Evaluating Individual Level Responses to Exercise for Health
456 Outcomes in Overweight or Obese Adults. *Front Physiol*. 2019;10. doi:10.3389/fphys.2019.01401

457 10. Huffman KM, Slentz CA, Bateman LA, et al. Exercise-induced changes in metabolic
458 intermediates, hormones, and inflammatory markers associated with improvements in insulin
459 sensitivity. *Diabetes Care*. 2011;34(1):174-176. doi:10.2337/dc10-0709

460 11. AbouAssi H, Slentz C a, Mikus CR, et al. The Effects of Aerobic, Resistance and Combination
461 Training on Insulin Sensitivity and secretion in Overweight Adults from STRRIDE AT/RT: A
462 Randomized Trial. *J Appl Physiol (1985)*. 2015;118(919):jap.00509.2014.
463 doi:10.1152/japplphysiol.00509.2014

464 12. Phillips BE, Kelly BM, Lilja M, et al. A practical and time-efficient high-intensity interval training
465 program modifies cardio-metabolic risk factors in adults with risk factors for type II diabetes.
466 *Front Endocrinol (Lausanne)*. 2017;8(SEP):1-11. doi:10.3389/fendo.2017.00229

467 13. Timmons JA, Atherton PJ, Larsson O, et al. A coding and non-coding transcriptomic perspective
468 on the genomics of human metabolic disease. *Nucleic Acids Res*. 2018;46(15):7772-7792.
469 doi:10.1093/nar/gky570

470 14. Álvarez C, Ramírez-Campillo R, Ramírez-Vélez R, Izquierdo M. Effects and prevalence of
471 nonresponders after 12 weeks of high-intensity interval or resistance training in women with
472 insulin resistance: a randomized trial. *J Appl Physiol*. 2017;122:985-996.
473 doi:10.1152/japplphysiol.01037.2016.-Our

474 15. Sears DD, Hsiao G, Hsiao A, et al. Mechanisms of human insulin resistance and
475 thiazolidinedione-mediated insulin sensitization. *Proc Natl Acad Sci U S A*. 2009;106(44):18745-
476 18750. doi:10.1073/pnas.0903032106

477 16. Ustinova M, Ansone L, Silamikelis I, et al. Whole-blood transcriptome profiling reveals signatures
478 of metformin and its therapeutic response. *PLoS One*. 2020;15(8 August):1-17.
479 doi:10.1371/journal.pone.0237400

480 17. Ross LM, Slentz CA, Zidek AM, et al. Effects of Amount, Intensity, and Mode of Exercise
481 Training on Insulin Resistance and Type 2 Diabetes Risk in the STRRIDE Randomized Trials.
482 *Front Physiol*. 2021;12(February):1-10. doi:10.3389/fphys.2021.626142

483 18. Tang Q, Li X, Song P, Xu L. Optimal cut-off values for the homeostasis model assessment of
484 insulin resistance (HOMA-IR) and pre-diabetes screening: Developments in research and
485 prospects for the future. *Drug Discov Ther*. 2015;9(6):380-385. doi:10.5582/DDT.2015.01207

486 19. Ferrannini E, Mari A. How to measure insulin sensitivity. *J Hypertens*. 1998;16(7):895-906.
487 doi:10.1097/00004872-199816070-00001

20. Manley SE, Stratton IM, Clark PM, Luzio SD. Comparison of II human insulin assays: Implications for clinical investigation and research. *Clin Chem*. 2007;53(5):922-932. doi:10.1373/clinchem.2006.077784

21. Miller WG, Thienpont LM, Van Uytfanghe K, et al. Toward standardization of insulin immunoassays. *Clin Chem*. 2009;55(5):1011-1018. doi:10.1373/clinchem.2008.118380

22. Thambisetty M, Metter EJ, Yang A, et al. Glucose intolerance, insulin resistance, and pathological features of Alzheimer disease in the Baltimore Longitudinal Study of Aging. *JAMA Neurol*. 2013;70(9):1167-1172. doi:10.1001/jamaneurol.2013.284

23. Rodríguez-Mañas L, Angulo J, Carnicero JA, El Assar M, García-García FJ, Sinclair AJ. Dual effects of insulin resistance on mortality and function in non-diabetic older adults: findings from the Toledo Study of Healthy Aging. *Geroscience*. 2022;44(2):1095-1108. doi:10.1007/s11357-021-00384-4

24. Clausen JO, Borch-Johnsen K, Ibsen H, et al. Insulin sensitivity index, acute insulin response, and glucose effectiveness in a population-based sample of 380 young healthy Caucasians: Analysis of the impact of gender, body fat, physical fitness, and life-style factors. *Journal of Clinical Investigation*. 1996;98(5):1195-1209. doi:10.1172/JCI118903

25. Zheng Y, Ceglarek U, Huang T, et al. Weight-loss diets and 2-y changes in circulating amino acids in 2 randomized intervention trials. *American Journal of Clinical Nutrition*. 2016;103(2):505-511. doi:10.3945/ajcn.115.117689

26. Shah SH, Bain JR, Muehlbauer MJ, et al. Association of a peripheral blood metabolic profile with coronary artery disease and risk of subsequent cardiovascular events. *Circ Cardiovasc Genet*. 2010;3(2):207-214. doi:10.1161/CIRCGENETICS.109.852814

27. Bloomgarden Z. Diabetes and branched-chain amino acids: What is the link? *J Diabetes*. 2018;10(5):350-352. doi:10.1111/1753-0407.12645

28. Newgard CB, An J, Bain JR, et al. A branched-chain amino acid-related metabolic signature that differentiates obese and lean humans and contributes to insulin resistance. *Cell Metab*. 2009;9(4):311-326. doi:10.1016/J.CMET.2009.02.002

29. Floegel A, Stefan N, Yu Z, et al. Identification of serum metabolites associated with risk of type 2 diabetes using a targeted metabolomic approach. *Diabetes*. 2013;62(2):639-648.

30. Shalaurova I, Connelly MA, Garvey WT, Otvos JD. Lipoprotein insulin resistance index: a lipoprotein particle-derived measure of insulin resistance. *Metab Syndr Relat Disord*. 2014;12(8):422-429. doi:10.1089/MET.2014.0050

31. Flores-Guerrero JL, Gruppen EG, Connelly MA, et al. A newly developed diabetes risk index, based on lipoprotein subfractions and branched chain amino acids, is associated with incident type 2 diabetes mellitus in the prevend cohort. *J Clin Med*. 2020;9(9):1-17. doi:10.3390/jcm9092781

32. Zanetti D, Stell L, Gustafsson S, et al. Plasma proteomic signatures of a direct measure of insulin sensitivity in two population cohorts. *Diabetologia*. 2023;66(9):1643-1654. doi:10.1007/s00125-023-05946-z

33. Wang TJ, Larson MG, Vasan RS, et al. Metabolite profiles and the risk of developing diabetes. *Nat Med*. 2011;17(4):448-453. doi:10.1038/NM.2307

528 34. White PJ, Lapworth AL, McGarrah RW, et al. Muscle-Liver Trafficking of BCAA-Derived
529 Nitrogen Underlies Obesity-Related Glycine Depletion. *Cell Rep*. 2020;33(6).
530 doi:10.1016/j.celrep.2020.108375

531 35. Zhou M, Shao J, Wu CY, et al. Targeting BCAA catabolism to treat obesity-associated insulin
532 resistance. *Diabetes*. 2019;68(9):1730-1746. doi:10.2337/db18-0927

533 36. Lin C, Sun Z, Mei Z, et al. The causal associations of circulating amino acids with blood pressure:
534 a Mendelian randomization study. *BMC Med*. 2022;20(1). doi:10.1186/s12916-022-02612-w

535 37. Biswas D, Duffley L, Pulinilkunnil T. Role of branched-chain amino acid–catabolizing enzymes in
536 intertissue signaling, metabolic remodeling, and energy homeostasis. *FASEB Journal*.
537 2019;33(8):8711-8731. doi:10.1096/fj.201802842RR

538 38. White PJ, McGarrah RW, Herman MA, Bain JR, Shah SH, Newgard CB. Insulin action, type 2
539 diabetes, and branched-chain amino acids: A two-way street. *Mol Metab*. 2021;52.
540 doi:10.1016/J.MOLMET.2021.101261

541 39. Jang C, Oh SF, Wada S, et al. A branched-chain amino acid metabolite drives vascular fatty acid
542 transport and causes insulin resistance. *Nat Med*. 2016;22(4):421-426. doi:10.1038/nm.4057

543 40. Lynch CJ, Adams SH. Branched-chain amino acids in metabolic signalling and insulin resistance.
544 *Nat Rev Endocrinol*. 2014;10(12):723-736. doi:10.1038/nrendo.2014.171

545 41. Timmons JA, Anighoro A, Brogan RJ, et al. A human-based multi-gene signature enables
546 quantitative drug repurposing for metabolic disease. *Elife*. 2022;11. doi:10.7554/eLife.68832

547 42. Soto M, Cai W, Konishi M, Kahn CR. Insulin signaling in the hippocampus and amygdala
548 regulates metabolism and neurobehavior. *Proc Natl Acad Sci U S A*. 2019;116(13):6379-6384.
549 doi:10.1073/pnas.1817391116

550 43. Chen W, Huang Q, Lazdon EK, et al. Loss of insulin signaling in astrocytes exacerbates
551 Alzheimer-like phenotypes in a 5xFAD mouse model. *Proceedings of the National Academy of*
552 *Sciences*. 2023;120(21):e2220684120. doi:10.1073/PNAS.2220684120

553 44. Mutch DM, Temanni MR, Henegar C, et al. Adipose gene expression prior to weight loss can
554 differentiate and weakly predict dietary responders. *PLoS One*. 2007;2(12).
555 doi:10.1371/journal.pone.0001344

556 45. Newton RL, Johnson WD, Larrivee S, et al. A Randomized Community-based Exercise Training
557 Trial in African American Men: Aerobic Plus Resistance Training and Insulin Sensitivity in
558 African American Men. *Med Sci Sports Exerc*. 2020;52(2):408-416.
559 doi:10.1249/MSS.0000000000002149

560 46. Glynn EL, Piner LW, Huffman KM, et al. Impact of combined resistance and aerobic exercise
561 training on branched-chain amino acid turnover, glycine metabolism and insulin sensitivity in
562 overweight humans. *Diabetologia*. 2015;58(10):2324-2335. doi:10.1007/s00125-015-3705-6

563 47. Lee S, Gulseth HL, Langleite TM, et al. Branched-chain amino acid metabolism, insulin
564 sensitivity and liver fat response to exercise training in sedentary dysglycaemic and
565 normoglycaemic men. *Diabetologia*. 2021;64(2):410-423. doi:10.1007/s00125-020-05296-0

566 48. Dimenna FJ, Arad AD. The acute vs. chronic effect of exercise on insulin sensitivity: nothing lasts
567 forever. *Cardiovasc Endocrinol Metab*. 2021;10(3):149-161.
568 doi:10.1097/XCE.0000000000000239

569 49. Zheng C, Liu Z. Vascular function, insulin action, and exercise: an intricate interplay. *Trends*
570 *Endocrinol Metab*. 2015;26(6):297-304. doi:10.1016/J.TEM.2015.02.002

571 50. King DS, Baldus PJ, Sharp RL, Kesl LD, Feltmeyer TL, Riddle MS. Time course for exercise-
572 induced alterations in insulin action and glucose tolerance in middle-aged people. *J Appl Physiol*
573 *(1985)*. 1995;78(1):17-22. doi:10.1152/JAPPL.1995.78.1.17

574 51. Slentz CA, Bateman LA, Willis LH, et al. Effects of exercise training alone vs a combined
575 exercise and nutritional lifestyle intervention on glucose homeostasis in prediabetic individuals: a
576 randomised controlled trial. *Diabetologia*. 2016;59(10):2088-2098. doi:10.1007/s00125-016-4051-
577 z

578 52. Slentz CA, Bateman LA, Willis LH, et al. Effects of aerobic vs. resistance training on visceral and
579 liver fat stores, liver enzymes, and insulin resistance by HOMA in overweight adults from
580 STRRIDE AT/RT. *Am J Physiol Endocrinol Metab*. 2011;301(5):1033-1039.
581 doi:10.1152/ajpendo.00291.2011

582 53. An P, Teran-Garcia M, Rice T, et al. Genome-wide linkage scans for prediabetes phenotypes in
583 response to 20 weeks of endurance exercise training in non-diabetic whites and blacks: The
584 HERITAGE Family Study. *Diabetologia*. 2005;48(6):1142-1149. doi:10.1007/s00125-005-1769-4

585 54. Bouchard C, Rankinen T, Timmons JA. Genomics and genetics in the biology of adaptation to
586 exercise. *Compr Physiol*. 2011;1(3):1603-1648. doi:10.1002/cphy.c100059

587 55. Phillips BE, Kelly BM, Lilja M, et al. A practical and time-efficient high-intensity interval training
588 program modifies cardio-metabolic risk factors in adults with risk factors for type II diabetes.
589 *Front Endocrinol (Lausanne)*. 2017;8(SEP). doi:10.3389/fendo.2017.00229

590 56. Marott SCW, Nordestgaard BG, Tybjaerg-Hansen A, Benn M. Causal associations in type 2
591 diabetes development. *Journal of Clinical Endocrinology and Metabolism*. 2019;104(4):1313-
592 1324. doi:10.1210/jc.2018-01648

593 57. Pavone F, Piironen J, Bürkner PC, Vehtari A. Using reference models in variable selection.
594 *Comput Stat*. 2023;38(1):349-371. doi:10.1007/s00180-022-01231-6

595 58. Piironen J, Paasiniemi M, Vehtari A. Projective inference in high-dimensional problems:
596 Prediction and feature selection. *Electron J Stat*. 2020;14(1):2155-2197. doi:10.1214/20-EJS1711

597 59. *IDF Strategic Plan*.; 2023.

598 60. Keller P, Vollaard NBJ, Gustafsson T, et al. A transcriptional map of the impact of endurance
599 exercise training on skeletal muscle phenotype. *J Appl Physiol*. 2011;110(1):46-59.
600 doi:10.1152/japplphysiol.00634.2010

601 61. Manning AK, Hivert MF, Scott RA, et al. A genome-wide approach accounting for body mass
602 index identifies genetic variants influencing fasting glycemic traits and insulin resistance. *Nat*
603 *Genet*. 2012;44(6):659-669. doi:10.1038/ng.2274

604 62. Crossland H, Smith K, Idris I, Phillips BE, Atherton PJ, Wilkinson DJ. Exploring mechanistic
605 links between extracellular branched-chain amino acids and muscle insulin resistance: an in vitro
606 approach. *Am J Physiol Cell Physiol*. 2020;319(6):C1151-C1157.
607 doi:10.1152/AJPCELL.00377.2020

608 63. Roth Flach RJ, Bollinger E, Reyes AR, et al. Small molecule branched-chain ketoacid
609 dehydrogenase kinase (BDK) inhibitors with opposing effects on BDK protein levels. *Nat*
610 *Commun*. 2023;14(1). doi:10.1038/S41467-023-40536-Y

611 64. Sitryawan A, Hawes JW, Harris RA, Shimomura Y, Jenkins AE, Hutson SM. A molecular model
612 of human branched-chain amino acid metabolism. *Am J Clin Nutr*. 1998;68(1):72-81.
613 doi:10.1093/AJCN/68.1.72

614 65. Shin AC, Fasshauer M, Filatova N, et al. Brain Insulin Lowers Circulating BCAA Levels by
615 Inducing Hepatic BCAA Catabolism. *Cell Metab*. 2014;20(5):898-909.
616 doi:10.1016/J.CMET.2014.09.003

617 66. Sayda MH, Abdul Aziz MH, Gharahdaghi N, et al. Caloric restriction improves glycaemic control
618 without reducing plasma branched-chain amino acids or keto-acids in obese men. *Sci Rep*.
619 2022;12(1). doi:10.1038/S41598-022-21814-Z

620 67. Pavone F, Piironen J, Bürkner PC, Vehtari A. Using reference models in variable selection.
621 *Comput Stat*. 2023;38(1):349-371. doi:10.1007/S00180-022-01231-6/FIGURES/9

622 68. Piironen J, Paasiniemi M, Vehtari A. Projective inference in high-dimensional problems:
623 Prediction and feature selection. *https://doi.org/101214/20-EJS1711*. 2020;14(1):2155-2197.
624 doi:10.1214/20-EJS1711

625 69. Tripathy D, Cobb JE, Gall W, et al. A novel insulin resistance index to monitor changes in insulin
626 sensitivity and glucose tolerance: The ACT NOW study. *Journal of Clinical Endocrinology and*
627 *Metabolism*. 2015;100(5):1855-1862. doi:10.1210/jc.2014-3824

628 70. Wallace TM, Levy JC, Matthews DR, Homa T. Use and Abuse of HOMA Modeling. *Diabetes*
629 *Care*. 2004;27(6):1487-1495.

630 71. Levy JC, Matthews DR, Hermans MP. Correct homeostasis model assessment (HOMA)
631 evaluation uses the computer program. *Diabetes Care*. 1998;21(12):2191-2192.
632 doi:10.2337/DIACARE.21.12.2191

633 72. Matsuda M, Defronzo RA. Insulin Sensitivity Indices Obtained From Comparison with the
634 euglycemic insulin clamp. *Diabetes Care*. 1999;22(9):1462-1470.

635 73. Utzschneider KM, Prigeon RL, Faulenbach M V., et al. Oral Disposition index predicts the
636 development of future diabetes above and beyond fasting and 2-h glucose levels. *Diabetes Care*.
637 2009;32(2):335-341. doi:10.2337/dc08-1478

638 74. Marott SCW, Nordestgaard BG, Tybjaerg-Hansen A, Benn M. Causal associations in type 2
639 diabetes development. *Journal of Clinical Endocrinology and Metabolism*. 2019;104(4):1313-
640 1324. doi:10.1210/jc.2018-01648

641 75. Hagströmer M, Oja P, Sjöström M. The International Physical Activity Questionnaire (IPAQ): a
642 study of concurrent and construct validity. *Public Health Nutr*. 2006;9(6):755-762.

643 76. Godel H, Graser T, Földi P, Pfaender P, Fürst P. Measurement of free amino acids in human
644 biological fluids by high-performance liquid chromatography. *J Chromatogr A*. 1984;297:49-61.
645 doi:https://doi.org/10.1016/S0021-9673(01)89028-2

646 77. Vesali RF, Klaude M, Rooyackers O, Wernerman J. Amino acid metabolism in leg muscle after an
647 endotoxin injection in healthy volunteers. *Am J Physiol Endocrinol Metab*. 2005;288:360-364.
648 doi:10.1152/ajpendo.00248.2004.-Decreased

649 78. Jeyarajah EJ, Cromwell WC, Otvos JD. Lipoprotein particle analysis by nuclear magnetic
650 resonance spectroscopy. *Clin Lab Med*. 2006;26(4):847-870. doi:10.1016/J.CLL.2006.07.006

651 79. Wolak-Dinsmore J, Gruppen EG, Shalaurova I, et al. A novel NMR-based assay to measure
652 circulating concentrations of branched-chain amino acids: Elevation in subjects with type 2
653 diabetes mellitus and association with carotid intima media thickness. *Clin Biochem*. 2018;54:92-
654 99. doi:10.1016/J.CLINBIOCHEM.2018.02.001

655 80. Bürkner PC. brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw*.
656 2017;80. doi:10.18637/jss.v080.i01

657
658

659 **Table 1.**

	MP		S2		S-PD	
	Pre	**Post**	**Pre**	**Post**	**Pre**	**Post**
N	**179**	**109**	**116**	**85**	**149**	**122**
Gender (f:m)	100:79	58:51	58:58	43:41	89:60	71:51
Age	38 (18)	38 (18)	50 (14.5)	50 (17)	59 (11)	59 (12)
Weight (kg)	91.5 (20.1)	91.6 (21)	88 (17.5)	85.5 (13.2)	86.2 (17.4)	83.7 (16.5)
BMI	31 (5.5)	30.6 (4.766)	30.4 (4.88)	30 (5)	30.2 (4.45)	29.4 (4.6)
SBP (mmHg)	125 (14)	121 (16)	119 (16)	121 (18)	126 (16)	122 (21)
DPB (mmHg)	78 (13)	77 (14)	80 (14)	80 (14)	75 (12)	72 (13)
MAP (mmHg)	94 (13)	90 (12)	93 (15)	80 (14)	91 (13)	89 (13)
Resting HR	70 (12)	68 (12)	74 (16)	NA	NA	NA
VO2 max (mL/min/kg)	27.5 (10)	30.2 (11)	26.7 (7.56)	30.7 (9.6)	23.8 (7.05)	26 (8.25)
Fasting glucose (mmol/L)	4.61 (0.43)	4.6 (0.4)	5.26 (0.65)	5.33 (0.77)	5.78 (0.78)	5.81 (0.74)
2-hour glucose (mmol/L)	6.67 (1.51)	6.32 (1.9)	NA	NA	7.66 (3.43)	7.22 (2.3)
Fasting insulin (imm)	84.4 (65.4)	78.7 (51.7)	55.7 (43.7)	44.3 (33.5)	236 (234)	195 (199)
Fasting insulin (ELISA)	57.7 (42.5)	50.9 (37.6)	30.7 (43.7)	36.6 (33.2)	49.6 (58.4)	40.3 (47.2)
HOMA2 IR	1.2 (0.9)	1.07 (0.8)	1.07 (1)	0.81 (0.71)	1.12 (1.3)	0.95 (1.07)
Sum BCAA (µmol/L)	424 (133)	438 (139)	424 (90.4)	465 (85.5)	439 (99.1)	429 (96.1)
Leucine (µmol/l)	120 (40.6)	126 (39)	153 (36.2)	171.7 (39.5)	148 (37.9)	148 (35.1)
Isoleucine (µmol/L)	62.89 (22.6)	63.8 (21.7)	66.8 (20.1)	65.6 (22.5)	62.8 (19.7)	59.8 (14.5)
Valine (µmol/l)	239 (70.2)	247 (75.5)	208 (51.8)	230 (50)	221 (50.2)	220 (49.5)
HDL (mmol/L)	1.02 (0.39)	1.04 (0.4)	0.9 (0.33)	1.12 (0.33)	1.06 (0.39)	1.11 (0.37)
LDL (mmol/L)	2.64 (0.99)	2.5 (0.96)	2.45 (0.87)	2.72 (0.75)	3.12 (0.96)	2.96 (0.82)
TAG (mmol/L)	1.17 (0.75)	1.09 (0.72)	1.4 (0.85)	1.25 (0.78)	1.3 (0.83)	1.18 (0.95)
ALA (µmol/L)	321 (107)	326 (127)	419 (110)	405 (150)	399 (98)	390 (123)

660

661

662 Table 1. **Demographics. Median and IQR.** All available samples were utilised for baseline
663 modelling. In each analysis, only subjects with complete data were used in the pre/post
664 modelling, resulting in ~ 20% fewer subjects than recruited or completing the intervention.

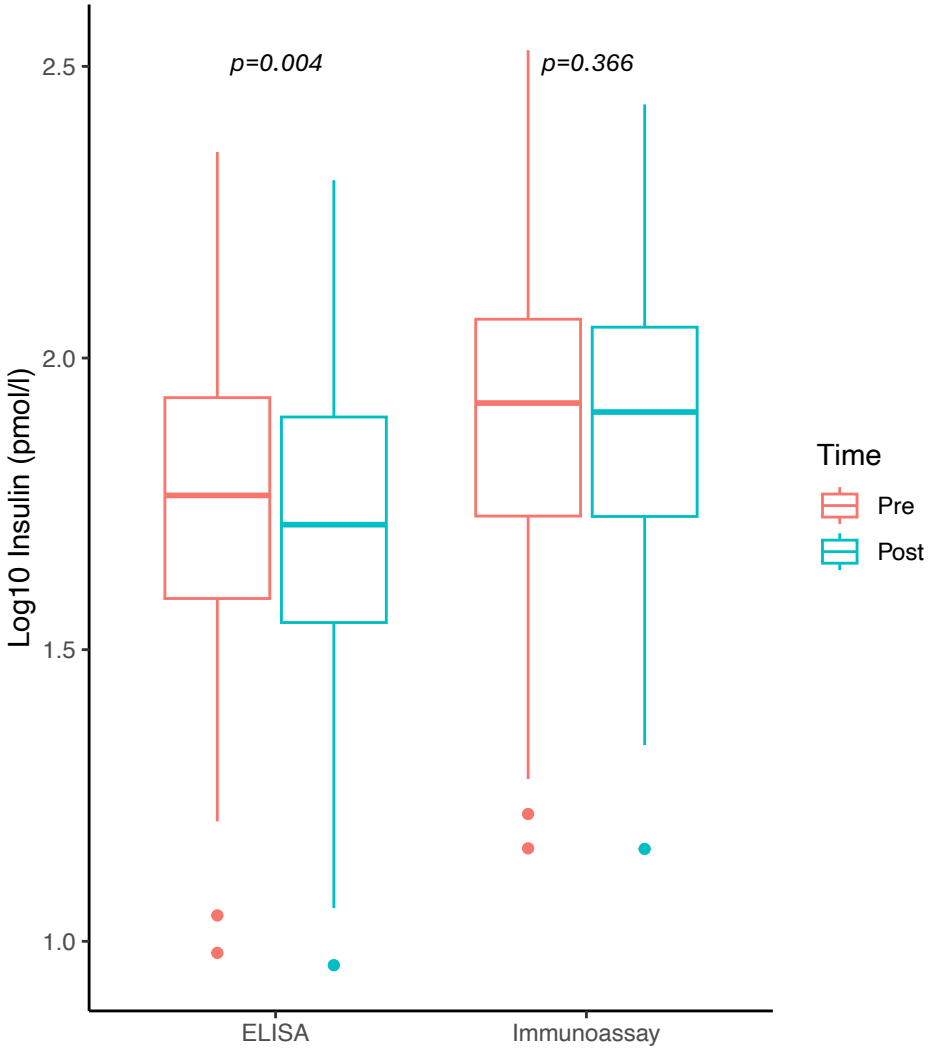
665

666 **Table 2**.

Cohort	Model	RMSE	MAE	r^2	Adj r^2	p value
MP n=179	1	0.208	0.167	0.272	0.255	<0.001
	2	0.205	0.164	0.302	0.282	<0.001
	3	0.208	0.161	0.309	0.281	<0.001
	4	0.206	0.161	0.327	0.295	<0.001
S-2 n=116	1	0.267	0.215	0.265	0.238	<0.001
	2	0.271	0.221	0.298	0.266	<0.001
	3	0.255	0.208	0.366	0.325	<0.001
	4	0.254	0.208	0.389	0.343	<0.001
S-PD n=149	1	0.295	0.237	0.202	0.18	<0.001
	2	0.281	0.227	0.265	0.239	<0.001
	3	0.293	0.229	0.304	0.269	<0.001
	4	0.291	0.223	0.339	0.301	<0.001

Model 1 = age+gender+$\log_{10}$BMI+fasting glucose
Model 2 = model 1 + sumBCAA
Model 3 = model 1 + HDL+LDL+TAG
Model 4 = model 1 + HDL+LDL+TAG+sumBCAA

667
668
669 **Table 2.** K-fold cross validation multiple linear regression predicting baseline $\log_{10}$ fasting insulin from
670 clinical and metabolomic variables.
671



672

673 **Figure 1**. Insulin assay statistical performance for identifying differences in pre vs post exercise
674 training samples (MP cohort). Fasting insulin (pmol/L) was measured using a high sensitivity
675 ELISA (left) and the Immulite 2000 automated analyser (right). P-values are calculated from
676 paired t-tests on log10 fasting insulin (pmol/L). Only the ELISA could detect a significant
677 treatment effect.

678
679

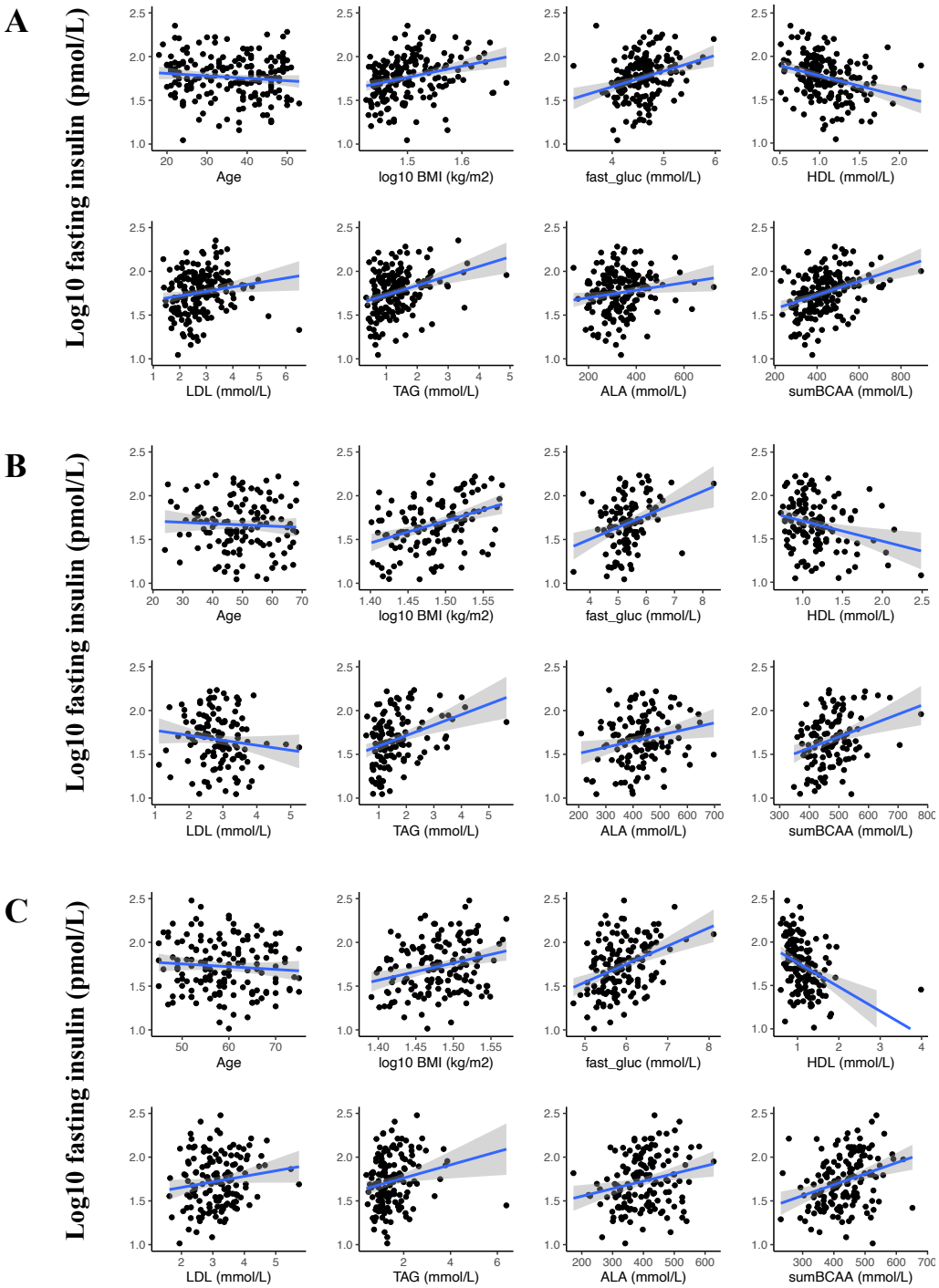


Figure 2. Scatter plots with OLS line of best fit (95% CI shaded) between log10 fasting insulin and clinical and metabolic variables. A) MP, B) S-2 and C) S-PD

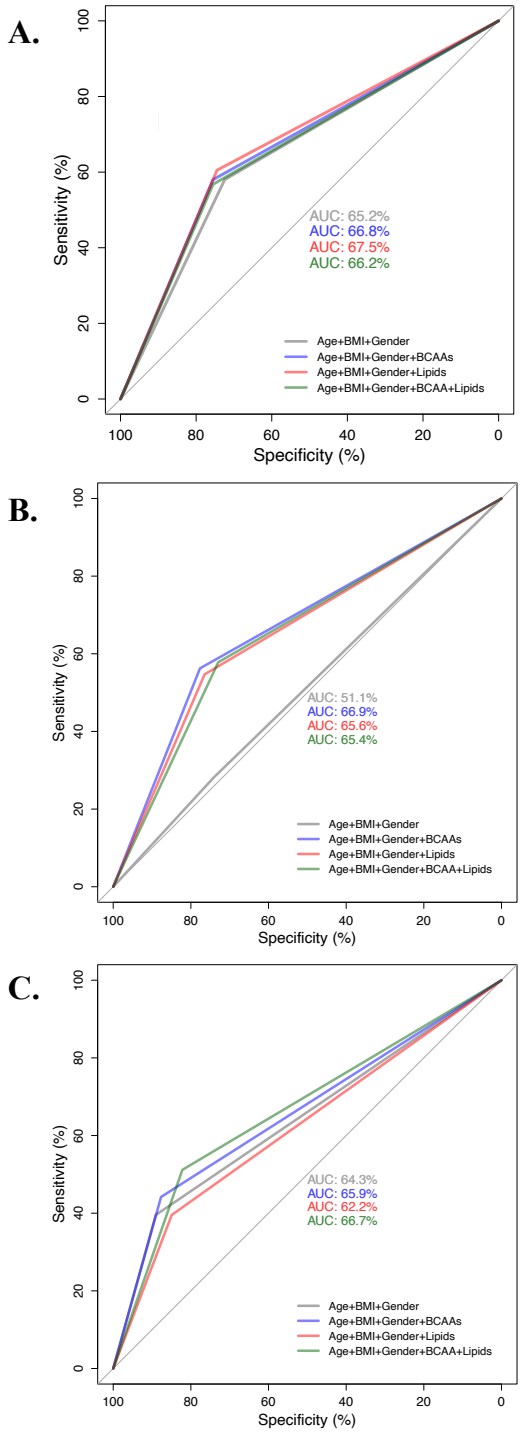


Figure 3. Utility of baseline clinical and metabolomic variables to classify HOMA2-IR status at 1.3 unit threshold reflecting the analysis of Marott et al. K-fold cross validation logistic regression ROC curves for A. MP, B. S-2 and C. S-PD.

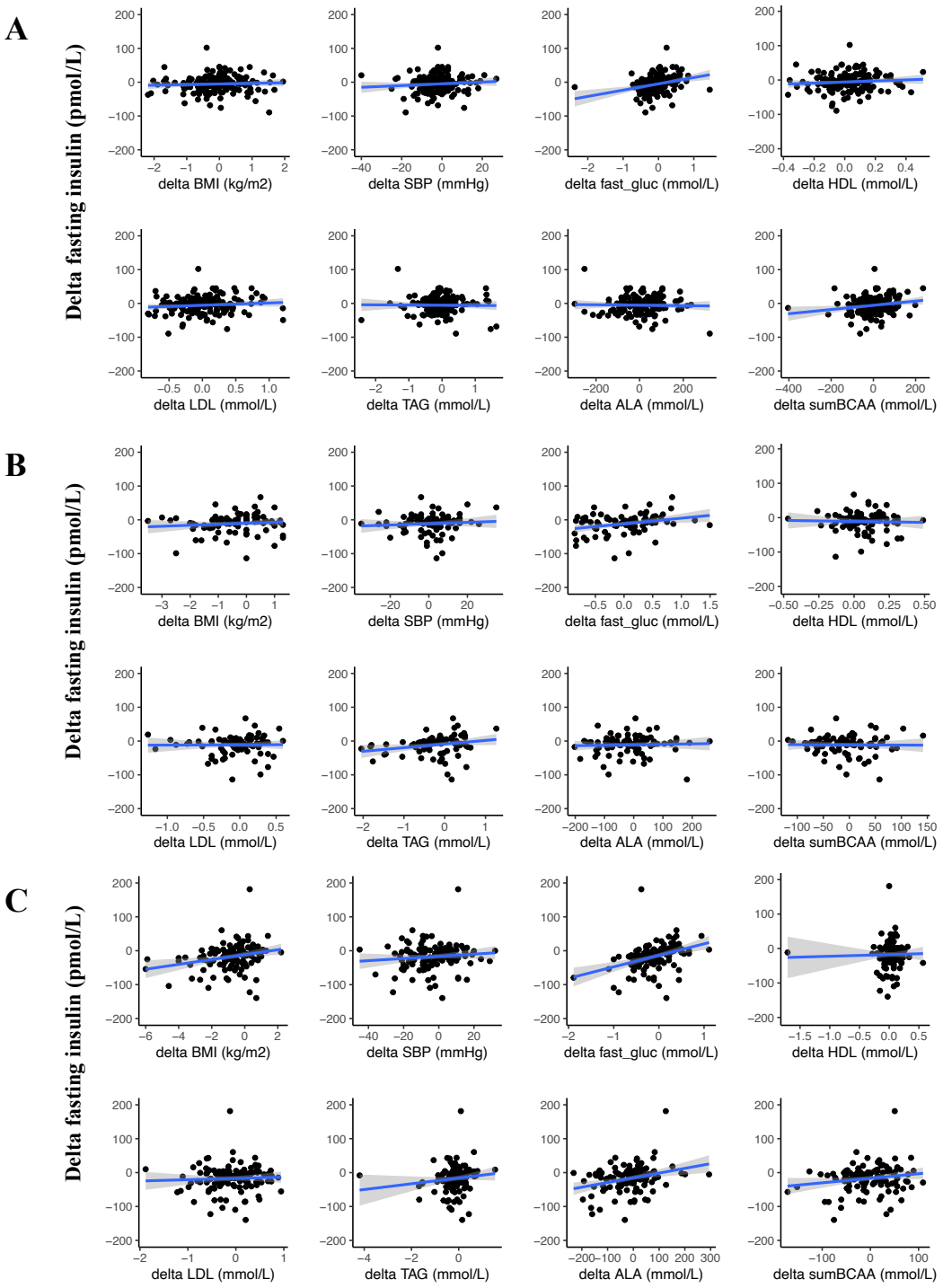


Figure 4. Scatter plots with OLS line of best fit (95% CI shaded) between delta fasting insulin and clinical and metabolic variables. A) MP, B) S-2 and C) S-PD

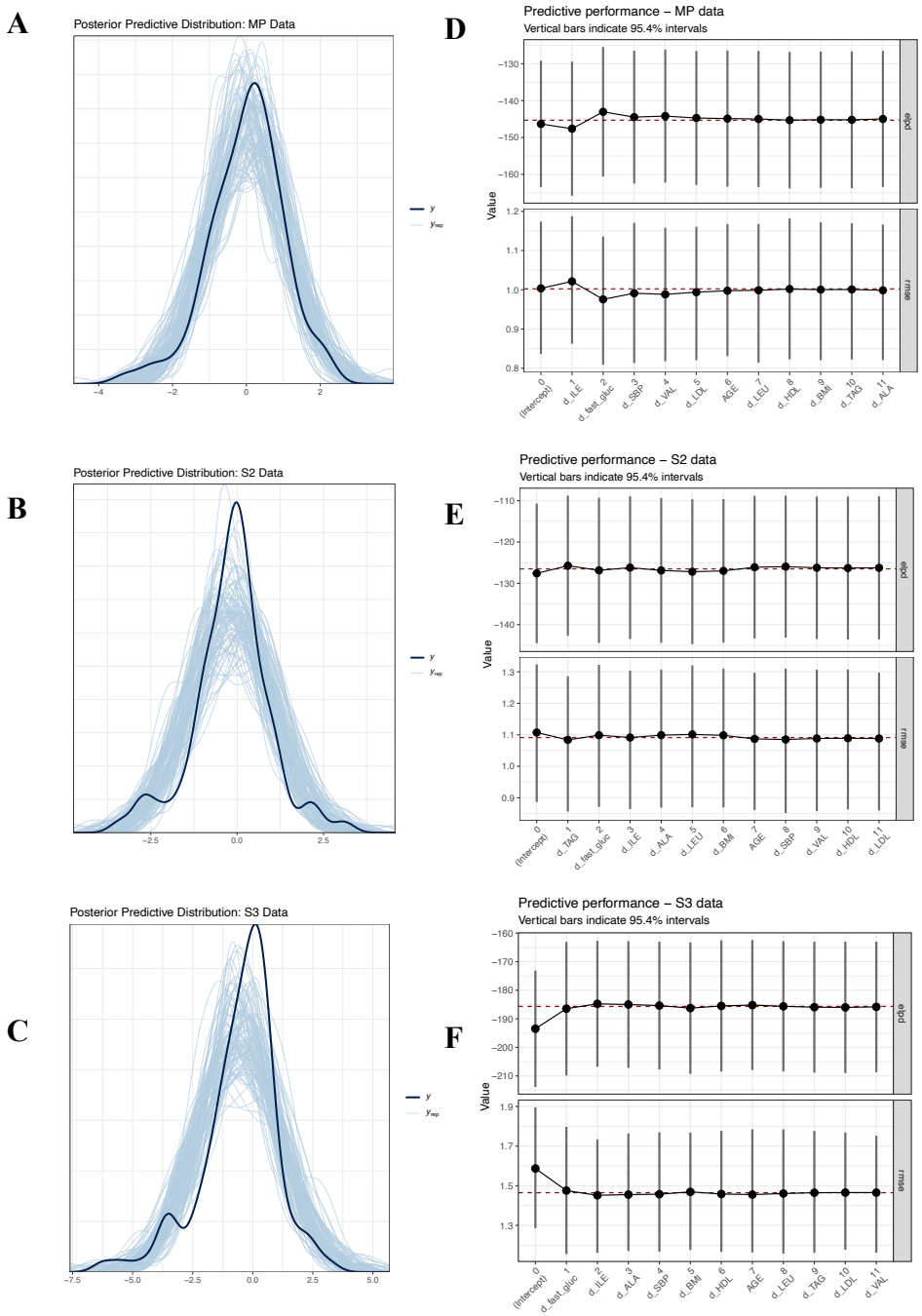


Figure 5. Left panel; Posterior predictive distributions for A. MP, B. S-2 and C. S-PD. Right panel; Expected log predictive density (Elpd, upper trace) and root mean square error (RMSE, lower trace) for D. MP, E. S-2 and F. S-PD. Elpd and RMSE are shown with 95% nominal coverage intervals.

699 **Table S1**. Linear regression modelling using log10 fasting insulin as dependant variable. Note the sample sizes reflecting all subjects
700 with data for each individual clinical parameter or metabolomic measure. In final integrated modelling only subjects with all
701 completes values were included.

702

META-PREDICT							
Model	**n**	**Intercept p**	**Intercept estimate**	**Estimates**	**95% CI**	**p**	**Adj r^2**
Age	188	<0.0001	1.8737	-0.0032	-0.0069 – 0.0005	0.0918	0.010
Gender	186	<0.0001	1.8254	-0.1227	-0.1928 – -0.0527	**0.0007**	0.056
BMI (log10)	188	0.5690	-0.2714	1.3532	0.7283 – 1.9781	**<0.0001**	0.084
SBP	175	<0.0001	1.1684	0.0048	0.0016 – 0.0080	**0.0038**	0.042
DBP	174	<0.0001	1.2888	0.0060	0.0020 – 0.0099	**0.0030**	0.044
MAP	174	<0.0001	1.1370	0.0066	0.0026 – 0.0106	**0.0013**	0.053
Fasting glucose	188	0.1037	-51.5267	25.7240	12.2694 – 39.1787	**0.0002**	0.066
HDL	185	<0.0001	2.0214	-0.2481	-0.3604 – -0.1358	**<0.0001**	0.089
LDL	184	<0.0001	1.5896	0.0599	0.0157 – 0.1041	**0.0082**	0.032
Triglycerides	187	<0.0001	1.6140	0.1110	0.0615 – 0.1604	**<0.0001**	0.091
Sum of BCAA	184	<0.0001	1.3981	0.0008	0.0005 – 0.0011	**<0.0001**	0.123
Isoleucine	184	<0.0001	1.4579	0.0045	0.0027 – 0.0063	**<0.0001**	0.115
Leucine	184	<0.0001	1.4408	0.0025	0.0015 – 0.0036	**<0.0001**	0.106
Valine	184	<0.0001	1.3818	0.0015	0.0009 – 0.0021	**<0.0001**	0.126

Alanine	184	<0.0001	1.5930	0.0005	0.0001 – 0.0009	**0.0148**	0.027
STRRIDE-2							
Model	**n**	**Intercept p**	**Intercept estimate**	**Estimates**	**95% CI**	**p**	**Adj r^2**
Age	128	<0.0001	1.7792	-0.0027	-0.0082 – 0.0027	0.3227	0
Gender	128	<0.0001	1.7628	-0.0886	-0.2063 – 0.0290	0.1386	0.009
BMI (log10)	128	0.0024	-2.8007	3.0012	1.7913 – 4.2110	**<0.0001**	0.154
SBP	111	<0.0001	1.2972	0.0029	-0.0017 – 0.0076	0.2139	0.005
DBP	111	0.0001	1.1314	0.0066	-0.0004 – 0.0136	0.0650	0.022
MAP	111	0.0002	1.1428	0.0055	-0.0009 – 0.0119	0.0911	0.017
Fasting glucose	118	0.0001	0.9260	0.1394	0.0582 – 0.2206	**0.0009**	0.083
HDL	128	**<0.0001**	1.8172	-0.2448	-0.4208 – -0.0687	**0.0068**	0.049
LDL	128	**<0.0001**	1.8172	-0.0606	-0.1486 – 0.0274	0.1753	0.007
Triglycerides	128	**<0.0001**	1.4320	0.1329	0.0677 – 0.1981	**0.0001**	0.107
Sum of BCAA	126	**<0.0001**	0.8578	0.0017	0.0009 – 0.0024	**0.0001**	0.117
Isoleucine	126	**<0.0001**	1.2102	0.0066	0.0029 – 0.0102	**0.0005**	0.085
Leucine	126	**<0.0001**	1.1699	0.0027	0.0008 – 0.0047	**0.0064**	0.051
Valine	126	**<0.0001**	0.8261	0.0035	0.0019 – 0.0050	**<0.0001**	0.13
Alanine	128	**<0.0001**	1.2359	0.0010	0.0003 – 0.0016	**0.0027**	0.062

STRRIDE PD							
Model	**n**	**Intercept p**	**Intercept estimate**	**Estimates**	**95% CI**	**p**	**Adj r^2**
Age	155	<0.0001	1.9753	-0.0046	-0.0116 – 0.0024	0.1944	0.005
Gender	155	<0.0001	1.7253	-0.0395	-0.1455 – 0.0665	0.463	0
BMI (log10)	155	0.1983	-1.2123	1.9683	0.7168 – 3.2198	**0.0023**	0.053
SBP	134	<0.0001	1.2980	0.0032	-0.0009 – 0.0074	0.1260	0.010
DPB	134	<0.0001	1.4408	0.0035	-0.0025 – 0.0096	0.2507	0.002
MAP	134	<0.0001	1.2897	0.0045	-0.0014 – 0.0105	0.1365	0.009
Fasting glucose	154	0.2343	0.3187	0.2357	0.1464 – 0.3251	**<0.0001**	0.146
HDL	153	<0.0001	2.0821	-0.3326	-0.4602 – -0.2051	**<0.0001**	0.144
LDL	151	<0.0001	1.5234	0.0568	-0.0180 – 0.1316	0.1355	0.008
Triglycerides	153	<0.0001	1.5573	0.0961	0.0311 – 0.1610	**0.0040**	0.047
Sum of BCAA	155	<0.0001	1.1048	0.0014	0.0007 – 0.0021	**0.0001**	0.091
Isoleucine	155	<0.0001	1.2519	0.0071	0.0038 – 0.0104	**<0.0001**	0.098
Leucine	155	<0.0001	1.3715	0.0022	0.0004 – 0.0041	**0.0159**	0.031
Valine	155	<0.0001	0.9665	0.0033	0.0021 – 0.0046	**<0.0001**	0.147
Alanine	155	<0.0001	1.4536	0.0006	-0.0000 – 0.0013	0.0675	0.015

703

704
705 **Table S2** - K-fold cross validation logistic regression model statistics for combined model predicting insulin resistance (HOMA2 IR
706 ≥1.3) using age, BMI, gender, sum of branched chain amino acids, HDL, LDL and triglycerides.
707

	MP	**S2**	**S-PD**
Accuracy (95% CI)	0.676 (0.602-0.744)	0.707 (0.615-0.788)	0.644 (0.562-0.721)
Kappa	0.339	0.347	0.27
Sensitivity	0.58	0.512	0.563
Specificity	0.755	0.822	0.706
PPV	0.662	0.629	0.59
NPV	0.685	0.741	0.682
Prevalence	0.453	0.371	0.43
Detection rate	0.263	0.19	0.242

708

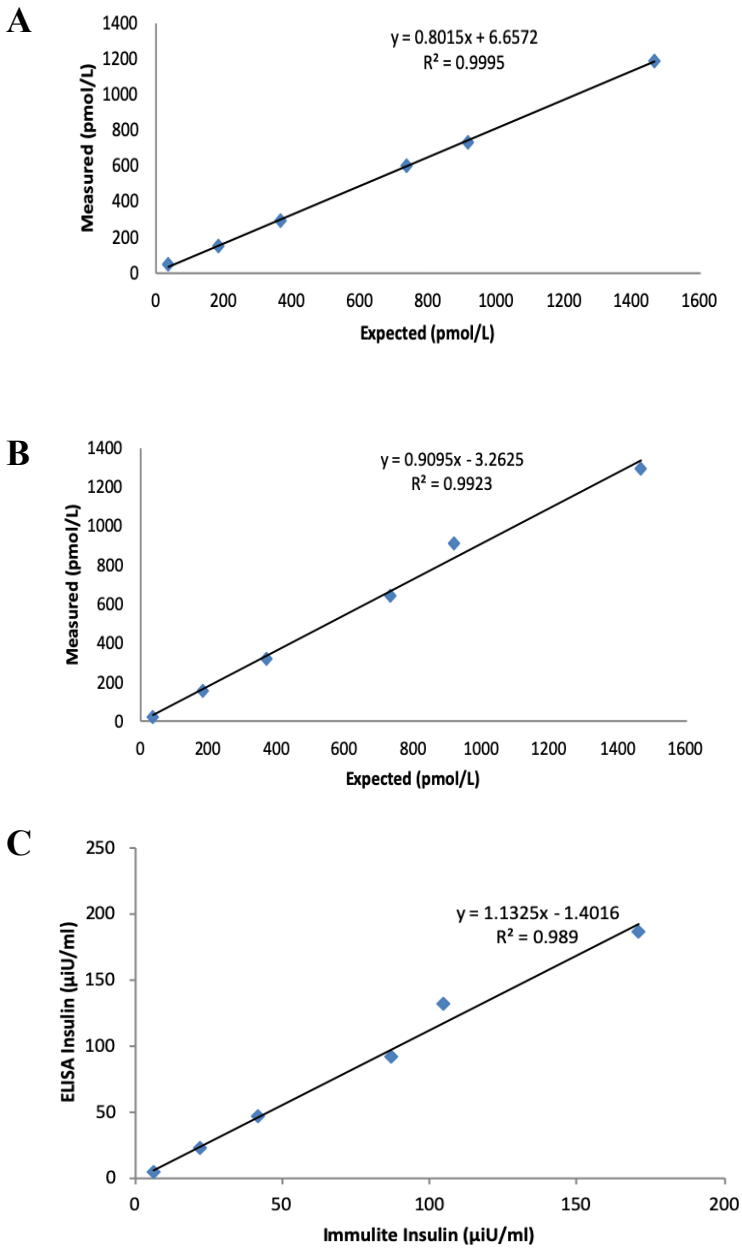
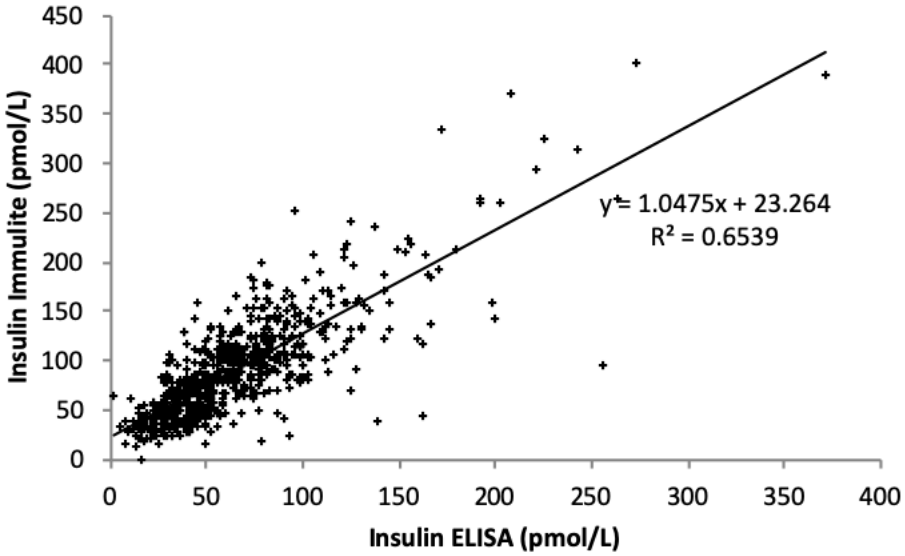


Figure S1. Performance of the insulin assays. A) (Immulite using titration international insulin standard), B) (ELISA using titration international insulin standard), and C) Immulite vs ELISA standard curve values using international insulin standard once converted to uIU/ml)



713
714 **Figure S2.** Correlation of participant fasting insulin measured using Dako ELISA and Immulite
715 2000 assays.
716
717

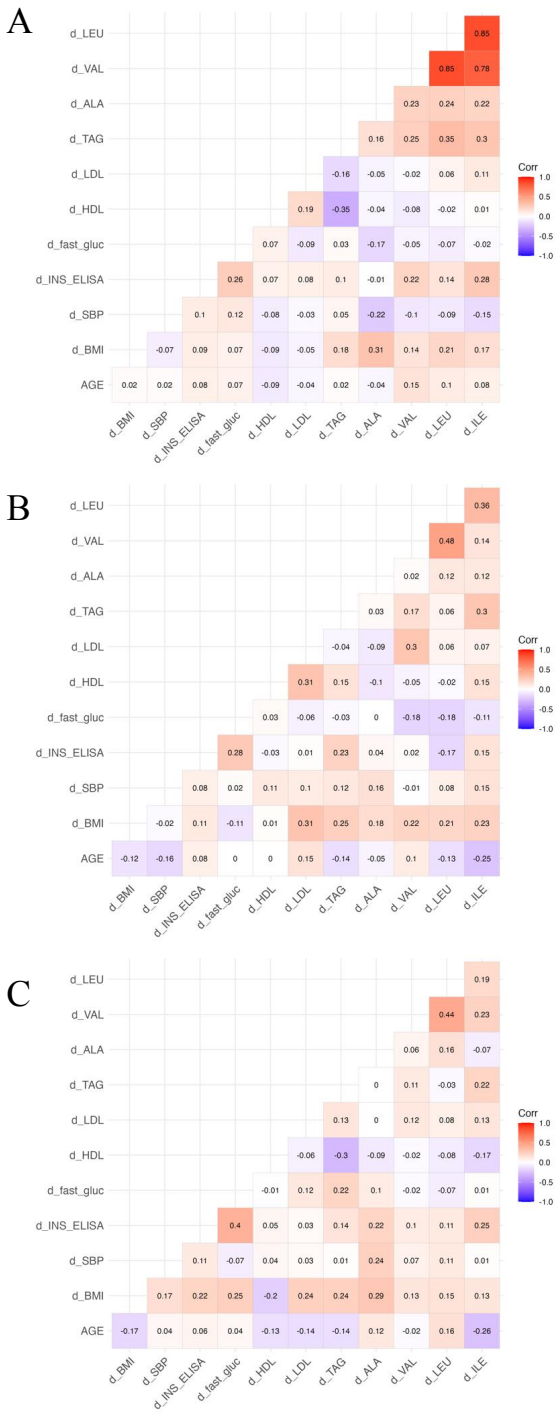


Figure S3. Correlation plots showing strength of correlation between the variables used in the generation of linear models to predict change in circulating insulin across lifestyle interventions. A. MP, B. S-2 and C. S-PD.

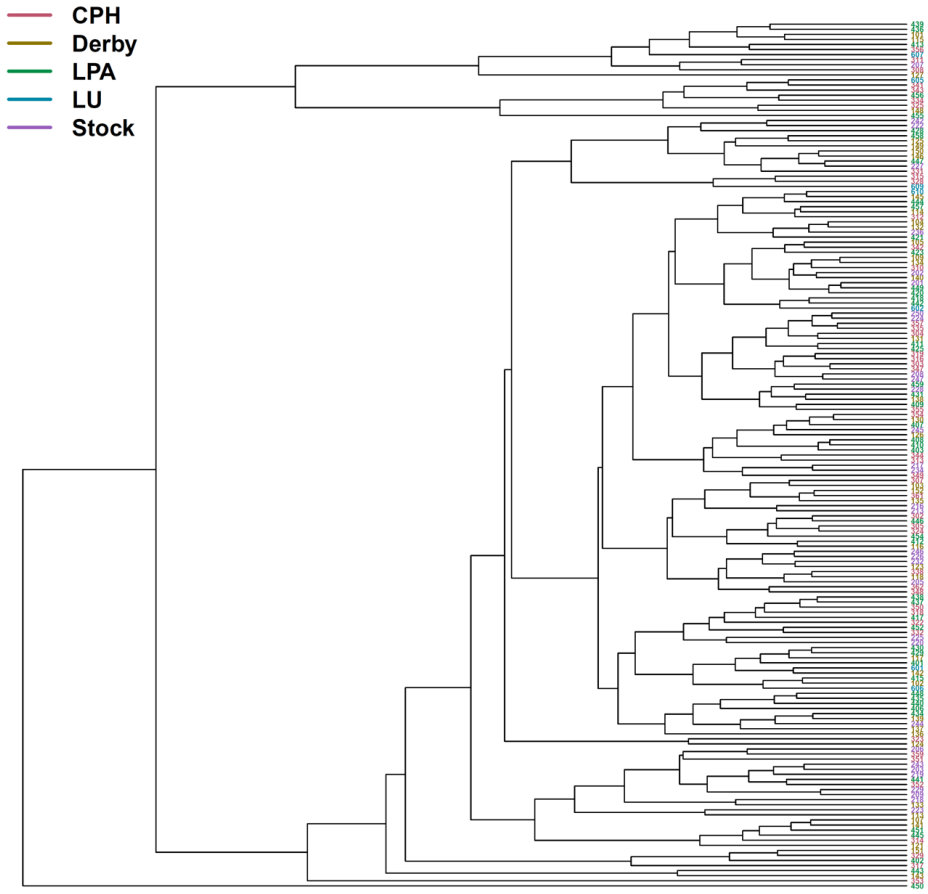


Figure S4. Hierarchical clustering of subjects by their metabolomic profile demonstrating that there is no centre specific bias.

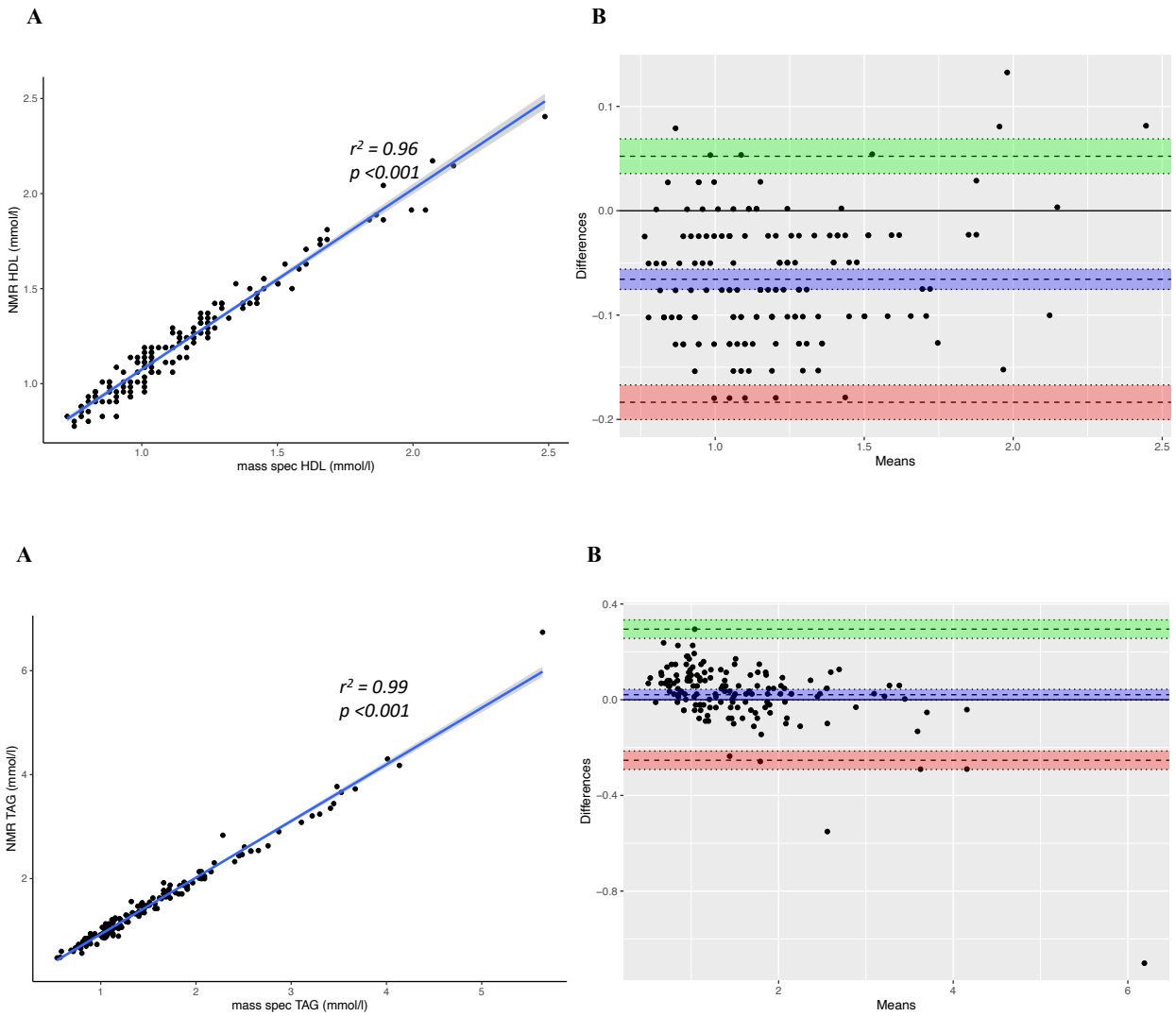


Figure S5-6. **S5)** S-2 HDL assay performance. A) Comparison of mass spectroscopy and nuclear magnetic resonance (NMR) measurement of HDL using linear regression. B) Bland-Altman plot to estimate of bias with 95% confidence interval shown in blue, with upper (green) and lower (red) limits. **S6)**. S-2 TAG assay performance. A) Comparison of mass spectroscopy and nuclear magnetic resonance (NMR) measurement of TAG using linear regression. B) Bland-Altman plot to estimate of bias with 95% confidence interval shown in blue, with upper (green) and lower (red) limits.

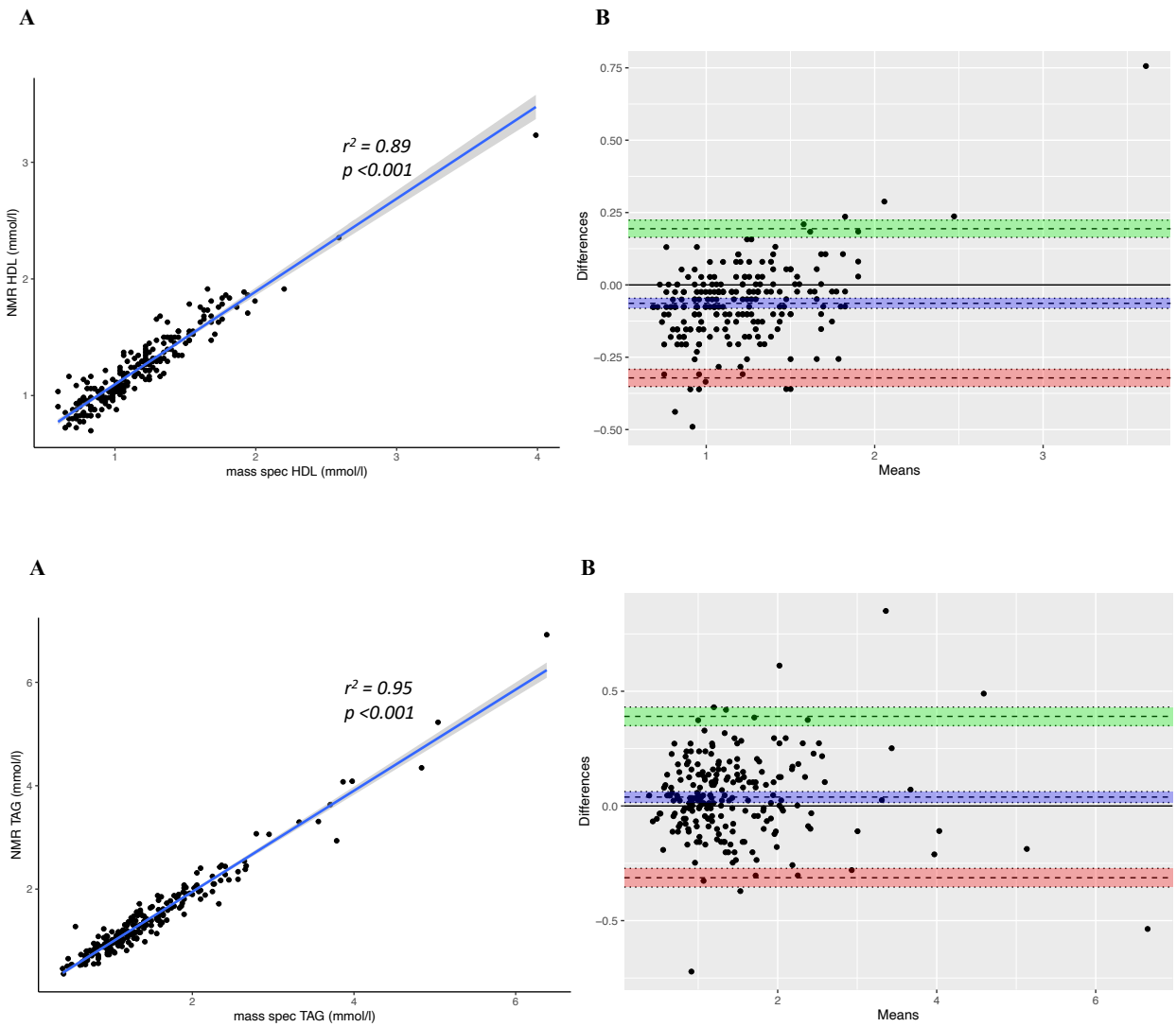


Figure S7-8. **S7)** S-PD HDL assay performance. A) Comparison of mass spectroscopy and nuclear magnetic resonance (NMR) measurement of HDL using linear regression. B) Bland-Altman plot to estimate of bias with 95% confidence interval shown in blue, with upper (green) and lower (red) limits. **S8)**. S-PD TAG assay performance. A) Comparison of mass spectroscopy and nuclear magnetic resonance (NMR) measurement of TAG using linear regression. B) Bland-Altman plot to estimate of bias with 95% confidence interval shown in blue, with upper (green) and lower (red) limits.